

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAACAGAAAACCTGTTAGAAATGTTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGTAAGTACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCCTTTTGTGCTGCACATGTAAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTTCTCTCTGAAATTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVICWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTTGCATTTTCTTGCAACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGG
TTTTTCATGGTGCCTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGA
CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
CATTCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTTCTGGATTGGTGTGTCAACTGCCCATACACTTACATGCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA
AACCATCAGGTTTCTGGGGAATGATAAAAAGTGTACCACCTCAGCATCAGGAAGTGAATACTTACT
CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTAATTTTCTTGGTT
ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT
GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTTGGTCCCAACACATTTCTTCATTCTGTTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCCTG
CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
AGAATACCGCACCATAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTG
ATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTA
GATATAAGAGGGGGGAAAAATGGAACAGGCGCTGACATTTTATAAACAAACAAAATGCTATGGTAGC
ATTTTTACCTTCATAGCATACTCCTTCCCCGTGAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
TGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTG
CCTTGAGATTGACTCATTAATAATCAGAGACTGTAACAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYPFWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGD
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALAEELS
RFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTD
PVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGMY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCTGGCTCTGAAGGGGTAGGCACGATGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTCGCTGCAGAAGAGCTTCCATCCAGGT
 GTCATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAGGCCAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAACTTGCAGCTATGGCTGGGTGGAGATGGATTTCGTGGTCATCTCTAG
 GATTAGCCCAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCTGATTGGAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCAATTCAGAA
 ATTATCACCACCAAGATCCCATATTCAACACTCAAACCTGCAACACAAACAAGAAATTTATGT
 CAGTGACAGTACCTACTCGGTGGCATCCCTTACTCTACAATACCTGCCCCCTACTACTCTCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTACAGAAGTTTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACATTTGTGAAATAAAGCAGCATTCAGAATGA
 AGCTGCTGGGTGGAGGTGCCACGGCTCTGCTAGTGCTTGCTCTCTCTCTTTGGTGTG
 CAGCTGGTCTTGGATTTTGTATGTCAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAAT
 CAGCAGAAGGAAATGATCGAAACCAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCATA
 TGAGGAATCAAAGAAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC
 GATGCTGGAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCCTT
 CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAAGAAAGTCCA
 CCCTTGGTTCCTAAGTGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT
 GCCCTTCTCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCT
 TTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTGCAAAGTGAAGGAC
 CTAAACATCTCATCAGTATCCAGTGGTAAAGGCCCTCCTGGCTGTCTGAGGCTAGGTGGGTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGACCTTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAAT
 GGCAGAAAGTTTAGCCCCGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAA
 GCTAAAATAAAGAAATAGAACAAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGTCTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAACTCTTATAAATTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTC
 AAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATCCTCAAAAAATTGCACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTGTGCTGAGACTAATCTTATTCATTTCTCTAATATGGCAACCATTATAACCTTAAT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTCCAACAAGAGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTTAGAAACTT

FIGURE 6

MARCFSLVLLLTISIWTRLLVQGSRLAEEL⁵IQVSCRIMGITLVSKKANQQ¹⁰LNFT¹⁵EAKEACRL²⁰LG
LSLAGKDQVETALKASFETCSY²⁵GWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIIITTKDPI³⁰ENTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKA³⁵AFKNEAAGFGGVP⁴⁰TALLVLALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQ⁴⁵QKEMIETKVVKEEKANDSNPNEESKKT⁵⁰DKNP⁵⁵EESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCCGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGTACCCGCGAGGAGGAGGCCACCTCAATGAGATGTTCCGCGAGGTTGAGAACTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACAGAC
TGGACAAATGGTCTTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAAGTGTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGACCGATGCCCTTGTGCCAGTGGCCCTCTGCCAGC
CCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA
CATCTTCTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTGTGCAATTTGTTACGT
CCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGTCTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTTGGCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAATG
CAACAAATGAATTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT
CTCCACTACCCACACAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAATAATTCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCAGCAAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
AAAACTCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATG
TTTTAGGTGTATGGACTGTTGCCACCATGTATTATCCAGAGTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTAAATTATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMCLTRDSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGLCFAFQRLFPVCTPLPVEGELCHDPASRLDLITWELEPDG
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAAALLGGEET

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCCACCTTGTGAACCTCTGCGTGGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCCT
AACTGGGTACTGGCCCTGGGCCAATGCGTCCCTGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCATTGGCATTGGAGCCCTCATCCTGACCCTTGTGCAGATAGCCCGGCTCATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAGGGTGGTCTGCTCTGGACAAAGTACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGGTCTGTCTTCTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCTAT
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCCTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
AAACCTCCGCTCTCTATTAATAAATAAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTGTCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAACAAAACAAACAAACAAA
AAGATTTTATTAAAGATATTTTGTTAACTC

FIGURE 10

RTGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLEFWTL
NWVLALGQCVLAFAGAFASYWAFHKPDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVLDKVTDLLLFFGKLLVVGGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCTCTTCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCTTCTTCTTCTTCTTTTACCCTGCTCATGTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACATCTGGTCTACTTCGGCGTGTGGG
CTCCTTCTCTTATCCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCGATTCCCGTGCTGGTACGCAGGCCTCTTCTTCTTACTCTCTC
TTCTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACTGAGCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGTCTCTGC
CCAAGGTCCAGGACGCCCAGCCAACTCGGCTCTGCTGCAGGCCTCGGTCATCACCTCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTG
TGGGCCTCATCATCTTCTCTGTGACCCCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGGCCGGGCTTTGACAACGAGCAGGACGGGCTCACCTACAGCTACTCCTTCTTCC
ACTTCTGCCTGGTGTGTCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTGAGGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCCTGAGCCGGG
CTTCTAGTCGTAGTGCCTTCAGGTCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCTTCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCACCACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTCTGGTACGTCCCCAGGGGACCCTGCCCCCTTCTCTG
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTCTAATAAACAAGCCAGTGCGGTGTAACAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFFLGLVLSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLGVRVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVWVKICASWAGLLLYLWTLVAPLLLRNRD
ES

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTTGTACCTTTGACCTCTTATTCTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTTCATTCATC
CTTGCCTGGATTGAGACGTGGTTTCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTTCCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTTAGAATATGAGTACTACTTTTGTAAATGTGAAAAACCTCACAGAAAGTC
ATCGAGGCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAATAAGGTGAAGAAAGTCTGTGCTGTATTCTTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC
AGAATGGAATTTTTTGTTCATGTCTCAGATTTATTTTGTATTTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTTCTGTTTTGCATTATGTGTATGGCCTGAAGTGTGGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVVRTFCLEVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQAGFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCCTGCCCCGCCCGGCTCCCTGCGCCGCCCGCCTCCCGGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGACAGGGCTGCCAT
CCGGCTGCCAGTGACGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACACGGTGCCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTTGCCGGCCTGCCGGCCTGCAGCTCCTGGACCTGTACAGAACAGATCGCCAGCC
TGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCGGGCACTGCCCCGCTGCGCCTGCCCCGCTGCTGCTGCTGGACCTCAGC
CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCAGAACCTCCACGACC
TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACG
CGCCTGCGGTGGCCGGCAACACCCGCTTGCAGCTGCGGGCCGAGGACCTGGCCGGCCTGGC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
TCTTCCCCCGCTGCGGTGCTGGCAGCTGCCCCGCAACCCCTTCAACTGCGTGTGCCCTGAGC
TGGTTTGGCCCTGGGTGCGCGAGAGCCAGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCGCCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGAATACGCCGACTTTGGCTGCCAG
CCACCACCACACAGCCACAGTGCCACCACGAGGCCCGTGGTGGGGAGGCCACAGCCTTGTCT
TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC
CACTGCCCCACGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA
ATGGGGGCACATGCCACCTGGGGACACGGCACACCTGGCGTGCTTGTGCCCCGAAGGCTTCAG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCACTACGCCGAG
GCCAACACGGTCCCTGACCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGTGAGTACAGGTCACCCA
GCTCGGGCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGGCCCGGGGCGGTGCCGGAGG
GCGAGGAGGCCTGCGGGGAGGCCCATACACCCCGAGCCGTCCACTCCAACACGCCCCAGTCACC
CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC
GCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGGGGCGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCGGAAGTGGAGGGAGTGAAGGTCCCCTTGGAG
CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
ACTCATGGGCTTCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
CTAACGTCCCCAGAACCAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCGGACCCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAGAACTGGAAAGGAAGATGC
TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT
GGGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAGACAAACGATGATATGAA
GGCCTTTTGTAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVP LLLP LLLLLLALGPGVQGCPSGCQCSQPQT V FCTARQGT TVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLRLY
LGKNRIRHIQPGAFDTLDRLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILD TANVE
ALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNSLSLQALPGDLSGLFPRLRLLLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALSSSLAPT WLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCFPSTCLNGGTCHLGRHHLACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRS LTLGIEPVSPTSRLRVGLQRYLQGS SVQLRSLRLTYRNLSGPDKRLVTLRLPASLA EY
TVTQLRP NATYSVCVMPLGPRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLP LLIAPALAAV
LLAALAAVGAAYCVRGRAMAAAAQDKGQVGP GAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFP GPGQLQSPLHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAAGTGTACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAAGTGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAAGAGCTGACTGAGGAAGGCTCTCCCAAGGACAGACTGCTCTTGG
CTTCTGTATGCCTCTGGACTTGGTGTAAATCAAGTCAGGCAAAGGCTCTTGTATATTATACAT
TTGGAGCTCTTGGGGCAATCTAATAGCCCATGTTTGGTAAGTAGACTTTAGTGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTTTTCAGCTTTCATGATC
CAGATTTGCTTGTTAAGACCAAATATTCAGTTGAACCTCCTTCAAATCTTGTAAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTTCTTTAAATGATTAG
TTTGGCTGATTGCCCCATAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTTGTAATGGTGGCTATAGAAAACAACATGAAATATTATACAATATTTTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTCTTTGCTAATTTGG
AAGATTAACCTCATTTTAAATAAAATTATGTCTAAGATTAAAAA
AAA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTTLTSDESVDHTTAGRVVAGQIFLDSESEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRQMQEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYL PQNIQAAREMFELTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTTFGALGGNLIHMLVLSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCCATTCCTGCAGTGGAATTCATGAAC TAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAAC TAAAGATTGAAGACATCACTG
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACTTTTGAT
AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGTCTCTGGGGTCTGATCAATAATGCTGGTGTT
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA
TTTTCTGGATACCTCTGTCTCAGATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCATTCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAATTGAAAGACTTGCCCATTCAAATGA
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKG FHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNOVGEGKLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDLSLRDMK
AFGVHVSCEPGLFKTNLADPVKVEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
CATTCAGAAATATATTCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCGGT
CGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACCTGCAGGAGCATTT
TTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATTCCCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTT
GCCAATCTGGGCATGTCTGAACAACGGGTTATAAACTGTATCAGGTTCCCTGTATGTCCACTGG
TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG
TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAGAGTATATGCAAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTGAGGAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGA
ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATTCTGAATTTCTTCATTTCATGTGTT
ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACCTACAACCACCATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
GAGCAGCCCAGAAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
CTACATTTTGATCCCTTTTAACTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
ATTTCTATTGTTTTTACTATGTTGAGCTACTTGACGTAAGTTCATTTGTTTTTACTATGTTTACC
TGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTCTTTAGAATTGGAAGAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
GAGACCATGTCTATTAAAAAATAAAATGGAAGCAAGAATAGCCTTATTTTCAAAATATGGAAA
GAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHNLNTSDTEGFLLEVKGEAKNSITDSQMDDVEVVYTIIDIKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIIITESCSTHRLHSYKPKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEEGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLDLDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTE

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGCGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGAGCGAGCGGACCGAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCGCCGCCGCCACACCCCTCTGCGGTCCCCGCGGCGCTGCCACCCCTTCCCTCCTTCCCC
GCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCGAGGTACCAGCCCGCGCCTCT
GCTTCCCTGGGCGCGCGCCCTCCACGCCCTCCTTCTCCCTGGCCGCGCCTGGCACCAGGGGACCGTTGCCTGA
CGCGAGGGCCAGCTCTACTTTTCCGCCCGCTCTCCTCCGCTGCTCGCTCTTCCACCAACTCCAACCTCTCTCCC
TCCAGCTCCACTCGTAGTCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCAA
GGTGGGAACGCGTCCGCCCGGCCGACCATGGCACGGTTCGGCTTGGCCGCGCTTCTCTGCACCCTGGCAGTGCTC
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCAGCGTCTTTACGTGTCCAAAGGCTTC
AACAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTGAAGATCTGTCCCGAGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGAATCATTTG
CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTG
AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAAATTTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
TTCGCGCTGGTGAACCTCCAGTACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGTCCCTCGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTGCTCAAGGC
TTAGCGGTGGCGGAGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGCTCTCGTGACTGTGAAGCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCCCTTTCAACATTGAATCGGTGATGGATCCCATGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCCAACACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCTCCCTTCGAGCAAC
GTTTGCACGATGAGAGGATGGCTGCAGGAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGCAGTGACAGGAAATGGAATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGTGGTGTCCGTCTGGGGCA
CAGGCTACCTCCTCACTGTCTTCTGCATCTTGTCTCTGTTATGCAGAGAGAGTGGAGATTAATTTCTCAAACCTCTGAG
AAAAAGTGTTCATCAAAAAGTTAAAAGGCACCAAGTTATCACTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTCTGACTTTGTTTTCTCATTGAGTTTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGGTTCCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTTATTTGTTTGTATGTTTTTTCTCATTTCGTTTGTGGGT
TTTTTTTCCAACGTGATCTCGCCTTGTTCCTTACAAGCAAACAGGGTCCCTTCTGGCACGTAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTATTAAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQGCQPPKPLPAGRISRISSESASFARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGAATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG
AGCCCCGAGAAGAAAATTCTATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGAGGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTGTTTGTTGCTCTCTCTAGTGTCTTCTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISL L L L L L L P L M L M S M V S S S L N P G V A R G H R D R G Q A S R R W L Q E G G Q E C E C K D W F L R A P R R K F M
T V S G L P K K Q C P C D H F K G N V K K T R H Q R H H R K P N K H S R A C Q Q F L K Q C Q L R S F A L P L

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CAACTTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTGTG
CTCTGTGCAGAGCCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTC
GAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTA
GCAGAATGAGAGAAGACATTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCTC
CCAATGTTGTCCCTTTCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAACGGTCAAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTPLPHLMALLGCWQPLCKSYFPYIMAVLTPKSNRKMESKKRELFSSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFEPWTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
CCACTGACGCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPAMPATCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAAGTT
 CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATCTAAGTGGAAATCATGTCGG
 GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGCTTTGGCCATGATGTTTACC
 TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT
 TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
 CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
 CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAAATTGACAGTTGAGCTTTTCCAAATCAC
 AAATAAAGCCATCAGCAGTGCTCCCTTCTGCTGTTCCAGCCACTGTGGACATTGGCCATCCTCA
 TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCAGGTT
 ATGGAAGGCGGCCAAGTGGAAATATAAGCCCCCTTCGGGCATTTCGGTACATGTGGTCGTACCATTT
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG
 TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
 TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG
 GATTCCGAGAATCATTGTGATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
 CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
 AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
 ATTCAAAATCTTGTCGAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
 TTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
 AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
 CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTTGATC
 TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
 AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
 GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
 TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
 TTAAAAGACCTAATAAACCTATTCTTCTCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYDNDLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
ILIFFVVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGT VVKGSFLISVVRIPRIIVMYMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQNA YTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FII FLGKVLVVCFTVFGGLMAFNYNRA FQVWAVPLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCTTAGA
ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
GAAAATTTTTTGAAGGAAATTTGCCCTTCTTCAAACAAGGGTGTCAATCTGATATTTATGAGGAC
TGTGTCTCTACTATGAAGGCATCTGTTATTGAAATGTCTTGTGTTTGTCTGGTGACTGGAGTAC
ATTCAAACAAGAAACGGCAAAGAAGATTAAAGGCCCAAGTTCAGTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGAAAGATCATCGATCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAAATACTTGTTCGGAAGGTTGTCTGGACAGTCT
GGTTACAAAGGGAGTTATTTCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACTACCCATCAGCTCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
CACCTTGCCAAGGCCATCCCTTCTGCTGCTTCTACCACCAGCATCCCAGACCACAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAGCAGCCAAAACAGGCCC
AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC
GGATGTGAGCTGGGACTTGTCCAAAAGAAGATTGAGCACACAGTCTTTGGAGCCAGTATCCC
TGGGAGATCCAACTGCAAAATTGACTTGTGCTTTTTTAATTGATGGGAGCACCAGCATTGGCAAA
CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCCAAGCTCTTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCTGCTACTCACTTTAACCTCAAGACAC
ACACGAATTTCTGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
GTAGGTGCGGCCATCTCCTTTGTGACCAAGAACTTCTTTTCCAAAGCCAATGGAACAGAAGCGG
GGCTCCCAATGTGGTGGTGGTATGGTGGATGGCTGGCCACGGACAAAGTGGAGGAGGCTTCAA
GACTTGCGAGAGAGTCAGGAATCAACATTTCTTCATCACCATTGAAGGTGCTGTGAAAATGAG
AAGCAGTATGTGGTGGAGCCCACTTTGCAAAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC
GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCTGCAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCTGGCCTGCAGCAAGACCTGCTTGAACCTGGCTGACATTGGCTTCGTCATCGAC
GGCTCCAGCAGTGTGGGGACGGGCAACTTCCGCACCGTCTCCAGTTTGTGACCAACCTCACCAA
AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCGGTGCAGTACACCTACGAACAGCGGC
TGGAGTTTGGGTTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGCACCCAGCACGGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA
GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCGACGGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCCTGGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACTCCTTCTTTGTGGACGA
GTTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
AGCCTCGGAACTGAATTTCAGAGCAGGCAGAGCACCAGCAAGTGCTGCTTTACTAACTGACGTGTT
GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
AAATGTCTTGTATTATTCTTTGCCATCATGCTTTTTTCATATCCAAAACCTGGAGTTACAAAGA
TGATCACAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTTACAT
TTTGACAATTGTTTCAAATAAATGTTTCGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
AGCTTTTGTGAGATTTTAAAGTTGTTATTCTGATTTGAACTCTGTAAACCTCAGCAAGTTTCAT
TTTTGTGATGACAATGTAGGAATTGCTGAATTAATGTTTAGAAGGATGAAAAATAAAAAAAAAA
AAA
AAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHTNLKTHNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLLQPLVKRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTLQFVTN
LTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARHDSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPYCSGLGYMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLKVNHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCCTTGCCCAAGAGAGTACACAGTCATTAAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTCTGCCCGGAAAGAGGGAAGTCGTGGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAAGTGCAGAGCTGCCGAAATGGCTCATGGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGGCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAGGGTCAGATTTTGTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTCACTCCAACTAAGATTTGTTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGTATGGAGACAACCGCATGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCCCTCATCCCTTGTTCATGACGGCACGTGCGTCCCTTGACAAGGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAAAGAAAGAAATGCTCAGACCTGGGGGCCAGTCAATGGGTACCAAGAAA
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATGGCACCGTGGTGTCTTTCTTTTGTAACTCCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTGGCCAGCAGAAATGGAGAGTGGTCAAGGAAACAGCCCATCTGCATAAAAGCTGCCGA
GAACCAAAGATTTAGACCTGGTGAGAGGAGAGTTCTTCCGATGCAGGTTCAAGGAGAGACACCATACACAGCATAT
ACTCAGCGCCCTTCAGCAAGCAGAACTGCAGAGTGGCCCTACCAAGAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGGCGGGCACCATCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCGTGGCAGGACCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAGGGAGC
GTGGTTCTTAGTCTCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACTGAAAGTTGTTTGGGGAAATTTACCGGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCGATCAGCACCGAGTCCAGCCCATCTGCCCTGCTGCCAGTGGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTGGCAGACGTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAAGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCTTCCCGGA
CGAGCATCTCTGAGCCAGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTTACCAAGGTGCTGCCCTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGCCCTGAAGTGTGATTTGGCTGTGAACCTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAGAAGACCATATACAAACCTCTCCA
CTCCACTGACCTGGTGGCTTCCCAACTTTTCAAGTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCCTTTGGAGCTCTCAAGTTCTAGAGAGCTGCCCTGTGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTT
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCTTCCCATCTCTTGTACACATTTTAAATAAATAAGGGTTGGCTTCT
GAAC TACAAA
AAA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGIFYCAECRAGWYGGDCMRGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII
KRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNSYVLSGNE
KRTCQONGEWSGKQPIKACREPKISDLVRRRVLPMQVQSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVVSVDSSLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLGLVSWSYDKTCSHRLSTAFTKVLPFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTAATC
 TGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
 TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGGCGCTTCTGTTGC
 TGGTCTTGCTTGGCTCAGTCTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTCTGTATTGAGAACTCTGTA
 AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
 TCACAGCCACGGCTCCCTCCCGAGAGGTTTCTGCAGCTGCCACCATCTCTTAATGACAGACGAGCTGGCCTAGACA
 ACCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTA
 GGGCAGCGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA
 GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGAAAATTTCTGAAAACACCACTGCCCTGAAG
 TCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCAGTGAAA
 GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGG
 TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
 ACACTACGCTGTGCGTCTCTCGGGCAGCCCTGCCAGGTGCTGTGGTGACTGTGATGCGTGAACAGAAAGTCCGCA
 GCAGGAACAATGGACAGGCCCCGGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
 CCGAGGAGCAGCTTGGAAATAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTCATCTTCAATGTGCTGGATGGCG
 GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
 GCCCAGAAAGTGGGCTCATCTGATTGAGGCCAGTGAAAGACGTGTTACCTCGTCGTGTCCCGCCAGGTTGCGGCAGC
 GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACA
 CTCCAAGCCCTCCATCCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAAGACCCCGTGAATCTCTCG
 GCATGACCGTCGAGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTGAGCCCGGAGGAG
 TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGTCGAACTGACAGAGGTGAGCC
 GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
 CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGAAGTGGTCCCCATCCT
 GGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC
 TGGGCTTCTGCATTGTAGGAGTTATGAAGAATACAATGGAAACAAACCTTTTTTCATCAAATCCATTGTTGAAGGAA
 CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAATGA
 TACATGCTTGCTTGGCAAGACTGCTGAAAGAAGCTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
 TTTTATAGAAATCAATGATGGGTGAGGAGAAACAGAAAAATCACAATAGGCTAAGAAGTTGAAACACTATATTATC
 TTGTGAGTTTTTATATTTAAAGAAAGAACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
 ACACCTCAGAAAATATGATTCAAAAAAATTAACCTACTAGTTTTTTTCAGTGTGGAGGATTCTCATTACTCTAC
 AACATTGTTTATATTTTTCTATTCAATAAAAAAGCCCTAAAACAACATAAATGATTGATTGTATACCCCACTGAATT
 CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTTAATTTACAGCT
 AAAATATTTTTTAAATGCAATGCTGAGAAACGTTGCTTTTCATCAAACAAGAATAAATATTTTTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSpanyIDNVGNLHFLYSELCKGASHYGLTKDRKRSQDGCPCDGCASLTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRSPDIFQEAGWNSNGSWSPG
PGRSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTGDILLNVVGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
TTCCACCTTTCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAGCAAAGGAGTTCATGGCTAATTTCC
ATAAGACCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC
TTTGAAGAGGTACAGGCAGAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
CTTTACAGAGGGTCGCCATCCTCGTTCCTCCACCGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
TAAAAAGTTTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAAATTGGG
ACTGCTTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCAGTGGGTACAGGTTACGTTACAGTGGATATTTTGG
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTACTGGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCATGACCC
TGGATCTTTTGGTGATGTTTGAAGAACTGATTCTTTGTTTGAATAATTTTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTGTATTTCT
TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTATTTAAATTTGAAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATACAATACTGTTATTTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTCAGGT
GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
ATTAACATAATAAATATGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLTFLHSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLLGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWGWGGE
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACAACAG
ACGGGACAACCTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCCCCTCCC
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGATAAAATGGCTGGTTCTTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

CONFIDENTIAL

FIGURE 44

MALSSQIWAACLLLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTCCACATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCCTGTCAGCAGAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG
AAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCCTCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEVDIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTGAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGTCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTtaggtctattgcttggttggaattctggaggctcctgtttgggct
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCGWTSCNGFSLVLVLLLVVNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCTT
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSVPVEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGCCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCAATTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
GTGCCCTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
CCTTGAGGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCACTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
CATAAACC CGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGGAGCGGGGAATCTGGGATT CAGGG
CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAATAAGCAAAGAGGGCAATCGCCTCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCACACTCCCTCCTTAA
AACACCACCCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
AA
AA

FIGURE 52

MKFQGPLACLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALQGQTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVF
GHSGAWETSGGHGIFGSQGGGLGGQGNPGGLGTPWVHGYPGNSAGSFGMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGSSNSGGGSGSQSGSSGSGSNGDNNNGSSSSGGS
SSGSSSSGSSSGGSSGGSSGSGNSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE
ARGSGESGIQGRGQGVSSNMREISKEGNRLGGSGDNYRGQSSWGSGGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACTGCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAAGTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTTCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCTATCCGGGAGCGGCGTCGACCCCTCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCCTTTCTCCGCA
GGGCCCAGGAAGTGCATCGGGCAGGCGTTCCCATGGCGGAGATGAAAGTGGTCTGCGGTTGAT
GCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCACTGACTTTCTGAC
CCATCCACCTGTTTTTTTGAGATTGTCATGAATAAACGGTGTGTCAAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPQPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGE GILLSSGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLD FIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLAHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTGTGTTTTGCACTTACCCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG
TACAGCCTTTCCTTCATACCATTTGCAAGGGATGCTGTGAAGAAGTGTTCGCGTGTGTCTTGC
ATAATTTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAGCAATGTGTGCTT
GTGATTCGAACATTTGAGGGTTACTTTTGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTCTGTATCTTGTGGAGTGAATCTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCCAGCAGGCCGCCAACAATGCTCTGTCTGTGCCTG
TACGTGCCGGTCAATCGGGGAAGCCAGACCGAGTTCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTATCCCTCCAGGAATTCTCCACCTACCGCCAGTGGAGCAGAAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGGCGAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTTTAAGATTTTGGACAAAAAAGATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCGGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAGCATGGATAAAAAACGGCACGATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAAATCCCCGAGATCATCTTACTGGAAGCATTCCACG
ATCTTTGATGTGGGTGAGAATTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGGCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTTGGCGGGCAATGGCATCAACGTCCTCAAAATTGCCCCGAATCAGCCATCAAAATTCATGGCCTATGAGCAG
ATCAAGCGCCTTGTGGTAGTGACCAAGGAGACTCTGAGGATTACGAGAGGCTTGTGGCAGGGTCTTGGCAGGGGCC
ATCGCCCGAGAGCAGCATACCCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGGTGGCCGCTTCTACAAGGCTATGTCCCAACATGCTG
GGCATCATCCCTATGCCGGCATCGACCTTGCAGTCTACGAGACGCTCAAGAAATGCTGGCTGCAGCACTATGCAGTG
AACAGCGCGGACCCCGGGTGTGTGTCTTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGTAC
CCCCGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAACATATCTTGCAGACCCAGGGGGGCTTGGGGTGTACAGGGGGCTGGCCCCCACTTCATGAAGGTTCATCCCA
GCTGTGAGCATCAGCTAGTGGTCTACGAGAACCTGAAGATCACCTTGGCGTGCAGTTCGCGGTGACGGGGGAGGGC
CGCCCGGCGAGTGGACTCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACCT
AAGCTGTCTCGAGCCAAGCTGTGAAACCTTAGACGACCCGAGGGAGGGTGGGAGAGCTGGCAGGCCAGGGCTT
GTCCTGCTGACCCAGCAGACCTCCTGTTGGTTCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCCAGCAGG
CTCCGGGCTCACATGTGTAAGGACAGGACATTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCT
TAGTTCTTCCATTTACCCCTTGCAGCCAGCTGTGGCCACGGCCCTGCCCTCTGCTGTGCCGTGCATCTCCCTGTGC
CCTCTTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAGCCACATCCCAACCCCTCGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCCTGACTTCCCAACCTACAGCATTGACGCCAATTGGC
TGTGAAGGAAGAGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG
CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT
GGCCTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAAGGGCGAAGGGCAGAGCGTTTGTGTCTTGGGAGGGAAGGAAAGGTGTTGGAGGCCCTTAATTATGG
ACTGTTGGGAAAAGGTTTGTCCAGAAGGACAAGCCGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTGCCAACCCAGCAGGGGCGCAGC
GGGACCAACCCCATCTTCACTTGTGTCACTGCTTGAACCTATTTATTTTGTATTATTTGAACAGATTATGTCCT
AACTATTTTATAGATTTGTTAATTAATAGCTTGTCAATTTCAAGTTCATTTTATTCATATTTATGTTTATGTT
GATGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGGAGGAGGGGGGCTTGGGCCGCTGCAGTCACATCT
GTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAGGCAGCAGCCTGGCTCCTTTCTTGGCAG
GTTGGGAAGGGCTTGGCCCCAGCCTTAGGATTTAGGGTTTGAAGTGGGGGCTGGAGAGAGAGGGAGGAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTCTTTATTTCACTCTTTCTGAATGTCAAGGCAG
TGAGGTGCCCTCTCACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCCTCCCTCCAGC
CTTCTGCTGCCCTTGTCTAACAATGCCGGCCAACTGGCGACCTCACGGTTGCACTTCCATTCCACCAGAAATGACCTGA
TGAGGAAATCTCAATAGGATGCAAGATCAATGCAAAAATTTATATATATGAACATATAACTGGAGTCTGCAAAAAG
CAAAATTAAGAAAGATTTGACGTTAGAAAGTTGTCAATTAAGCAGCCTTCTAATAAAGTTGTTTCAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEATQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKRLRVFKILDKKNDGRIDAQEIIMQSLRDLGVKISEQQAEEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTGCGAGCAGGATGAAATGTTTCAGAGGCCGACAGCAGTGTTTGTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACCTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
GTGCCTTGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTCAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTGTATATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTGTAAATGGGTCAAATGATTCACTTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIILLAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKNVQLTDAGTYKC
YIITSKKGKNANLEYKTGAFSMPEVNVNDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCCGGGGGCCCTATGGTCAGCCACCTCCAAGTTCTTACGGTGCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATGACCCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGTGGAGAGTCTGCATCATAGCCACCA
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCCTTGTAAATGG
AGTTAGTGTCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACAGTGGCTTGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTTGCATTTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC
AACTTGCCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCTGGGGGA
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGAAACCCAACATTGGGAGTTAGAGTCTGCTCCCGCCCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTAAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRRLPRNSNLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSGWNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFFNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCTTCTTGGCCC
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCGACCCCTGTCTTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGPLPLVLLLTLLGSSHGTGPGMTLQKLKESFLTNSSYESSFLELLEKLCLLLHLPSTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC
CAGGTGCCCCGTGCGAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLRLRWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIVVS
LLAALLLAVGLALLVRKLRKQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTGAGAGGTTTTGTTTTCTCTAGTTCTGTGCCTGCTGCACCAG
 TCAAATACTTCCCTTCAATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCTATTGTTATAGATCCTAGTGTGCCAGAA
 GATGAAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
 TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAACATGAA
 AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCATACCCAAGCAGTTTACA
 GAATGTGGAGAGAAAAGGCGAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAAAAATGAATATGGACCA
 CCAGGCAAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGTTGATGAGTACAATGAAGATCAGCCTTTC
 TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
 CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAATGTATGGAAAAGATTGTCAATTCCTT
 CCTGATAAAGTACAACAGAAAAAGCATCCATAATGTTTATGCAAAGTATTGATTCTGTTGTTGAATTTTGTAAACGAA
 AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA
 AGAATGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGTTTCACTTGTAGTACTGCCACTATTGTA
 AATAAGCTAATCCAAATAAAAAGCAGTGATGAAAGAAACACATCATGGCAGGATTACCTACATATCCTCTGGGAGGA
 ACTTCCATCTGCTCTGGAATTAATATGCATTTCAGGTGATTGGAGAGCTACATTCCCAACTCGATGGATCCGAAGTA
 CTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGATTGATGAAGTGAACAAAAGTGGGGCCATTGTTTAT
 TTTATTTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTT
 TCAGATGAAGCTCAGAACAATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAATACTGATCTCTCCAGAG
 TCCCTTCAGCTCGAAAGTAAGGGATTAACTGTAATAGTAAATGCCTGGATGAACGACACTGTCTATAATTGATAGTACA
 GTGGGAAAGGACACGTTCTTCTCATCATGGAACAGTCTGCCTCCAGTATTCTCTCTGGGATCCAGTGGGAACA
 ATAATGGAAAATTTACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCAGGAATGCAAGGTGGGCACT
 TGGGCATACAATCTTCAAGCCAAAGCGAACCAGAAACATTAACTATTACAGTAACCTTCTCGAGCAGCAATTTCTTCT
 GTGCCTCCAATCAGTGAATGCTAAATGAATAAGGACGTAAACAGTTTCCCAAGCCCAATGATTGTTTACGCAGAA
 ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCATTGAATCAGAGATGGACATACAGAAGTT
 TTGGAACCTTTGGATAATGGTGCAGGCGCTGATTCTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
 ACAGAAAATGGCAGATATAGCTTAAAAGTTCCGGGCTCATGGAGGAGCAAACTGCCAGGCTAAAATTACGGCCTCCA
 CTGAATAGAGCCCGTACATACCAAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAAACCCGCCAAGACCTGAAATTGAT
 GAGGATACTCAGACCACCTTGGAGGATTTAGCCGAACAGCATCCGGAGGTGCATTTGTGGTATCACAAGTCCCAAGC
 CTTCCTTGGCTGACCAATACCCCAAGTCAAATCAGACCTTGATGCCACAGTTTATGAGGATAAGATTATTCTT
 ACATGGACAGCACCAGGAGATAATTTGATGTTGGAAAAGTTCAACGTATATCATAAGAATAAGTGAAGTATTCTT
 GATCTAAGAGACAGTTTGTGATGCTCTTCAAGTAAATACTACTGATCTGTCCACAAAGGAGGCCAACTCCAAGGAA
 AGCTTTGCATTTAAACCAGAAAAATCTCAGAAGAAAATGCAACCCACATATTTATTGCCATTAAAGTATAGATAAA
 AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACCTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGAT
 CCTACACCTACTCCTACTCCTACTCCTGATAAAAGTCATAATCTGGAGTTAATATTCTACGCTGGTATTG
 TCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTAAGTACCACCTTTGAACCTTAACGAAGAAAAAATCTTC
 AAGTAGACCTAGAAGAGAGTTTAAAAAACAACAAATGTAAGTAAAGGATATTCTGAATCTTAAATTCATCCCAT
 GTGTGATCATAACTCATAAAAATAATTTAAGATGTCGGAAAAGGATACTTTGATTAAATAAAACACTCATGGATA
 TGTAAAACTGTCAAGATTAAATTTAATAGTTTCATTTATTTGTTATTTTATTGTAAGAAATAGTGATGAACAAAG
 ATCCTTTTTTCATACTGATACCTGGTTGTATATTATTGATGCAACAGTTTTCTGAAATGATATTTCAAATTGCATCAA
 GAAATTTAAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAAAATAACAACATTGGAATAAAAAA
 AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
RVYKCGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLVLDKSGSMGGKDRNLN
MNQAAKHFLQLTVENGWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVKGDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSEKNDGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRPPLNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTLDFSRASGGAFVVSQVPSLPLPDQYPPSQITDLDTVHEDKIIILTWTAPGDN
FDVGKVQRYIIIRISASILDLRDSFDDALQVNTDLSPKANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSANIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETLNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMVKNEKLNMAN
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY
NPGIYGHHDVAPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDDLKGGGEDVHLYR
KYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPPGGSEHQINSEAT
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC
TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC
AGAGAAGTGGAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACAGTGGAGTCCTTCGTCCC
AGGGCCCCCTCGCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAATCGTGATTA
ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTCTCATCAGGATATGAGTTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
AAAACACGGAAGGTACTTCTCTACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTGAGAACCCTGACATTTGTGCGGGGCTGAAGAGCAGGA
GCTCAGTTTGAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCCCGAAACGTTACAGTACTCATACCCCCTCAGCTCCAAGACTTAGACCCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGGCCGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA
TCCCCAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA
TGTGCAGATGGAAAATGATGCCAACACTTCCTTTTGCTTTTGTTCCTGTGCAACAAGTGAG
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT
CTGTGAGAATTACTTATTTCTTTTCTTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
GGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSIQIGPPEVALTTDEKSSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFDRFFVPAEK
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPDKTVIEYDYVRTTDICAGPEEQELSLQEEVSTQGTLLSQ
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRCLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites:

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCAAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGCGCCGGCCCGTGGACCTTCACCCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCCTCAGTCCCCTGTCAGTTCTCATCCTCGGCCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
CCCCCGCCCTGGTCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAACCTGCTGTA
TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTCATGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTATGGTCCCTGTTGGACTCTGTGCTTCCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
CCTTGGGATTTCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
TTGTGGAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGATGGCAGCCATCAGGGAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDNATSILQQPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTTPASLWKPPSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGGTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTCATCCTTGAGGGCCTCCTGGGATTC
ATTCTGTGTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCAGTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCGCTCCAATACTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCCCTACAGCCTGACAGGGTATGTGGAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCTCAACTTGAAACCCCATTCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAAGTGAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTT
CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCTCT
GCCACCGCTGCTTCCTGGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG
ACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCTCTCCTCCTTGCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCT
GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGGCCCCACAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCCTGCCCAACCAGTGTGTCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAACTGCCCCGAACCAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG
AGAGGGCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACCTGTCAAGATCGTCCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCAGAGG
ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTC
CTCGTCCACACATCGGTATCCCAAGCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCACAAGGTCCT
GGAACGTCTTCTAGCCAGACCCTGGAGCTGAAGGTACGGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTATTATATATTAATAAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARAPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGKTYT
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEPCRHPKRVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTAATCCATCCGTCACCTCTCCTGTCA
TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
ACAGGTGTCAGCACTGGGCTCAGTTCTCTCATTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCACTCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTGTGTC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGAATACTCTGCTGTGGCCTA
TTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCCAGTGGAATAATCCAGGCGGAAGTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT
ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAATAGGGGTCTTCTGGACTA
TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCTGACATGTCGGTTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAACCTCCCATAGTCATCTGC
CCAGTCAACCAGGAATCAGAGAAAGAGGCTCTTGGCAAAGGGCCTCTGCAATCCAGAGACAAGCAACAG
TGAGTCCTCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGG
CTGACATTACATTTAGTTTGTCTCTCACTCCATCTGGCTAAGTGTCTTGAAATACCACCTCTCAGGTGAAG
AACCCTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAA

FIGURE 84

MALMLSLVLSLLKLKSGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSSVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGDLSTDSTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT
LNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPESTSSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACAATGCTGCTGCTGCTGCTGCCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
GTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCCCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCTGTGCCCCCTCACAGCCCTCAAACCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTACCTGCAGAGCTCAGAACCTTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTT
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAACTGCAGAGACTCACCCCTGATTGAGGGATCACAGCCCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPDPSTTRS
SVLTLPQPQDHGTS LTCQVT FPGASVT TNKTVHLNVSYPQNLMTVFQGDGT VSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVS LQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGD TGIE
DANAVRG SASQG PLTEPWAEDSPPDQPPPASARSSVGE GELQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGGAATTCACCTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTGGATTGGAGTGGATAT
GGAACCTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIPIVVYDFGDAQKTASYSPYQGREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCTCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCTTCTCGATACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTATAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALEFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

90

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATATGCAACTGAC
TCGCTGCTGCTTCGTGTTCTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCTCCCCGGCTCAGAGGACCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCGAACCACAGCCCCC
CACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCGGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
CATAGTGATACCCCTACTACCCATCTGGGTCGACCCGGGGAGGCCACAGAGGCCAGGCCAGGGC
TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGTTGGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCTGTGGGCAGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTGTCTGTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
CGCGTGCTGTGTGTCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAGACCGGGGTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGCGAGCATGTGCTGGATCTGTTC
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MLTRCCFVFLVQGSLLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT
GGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGAAAAAGTGGGGCATCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLPGPTVGIIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKKGILLIVLLT.
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSCLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCACAAAACCTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAAAC
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTACACAACCTTGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTTGTTCCTGGGAGGCATCCTGCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCAGCAGGCCGCCTCCCAACTCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAATATTCTTGAAATTTAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGAAAACCTGAAAAA
AA
AA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPFI FVTQLGAQG TILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGS LPAGGAGVNPATQGTPAGR LPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTGACCATG
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCAGGCTCTCCCCAAGGGCCAGCCTGC
 AGAGCTGTCTGTGGAAGTTCCAGAAAACATATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC
 CGCTGCCCCGTGAGGGGGGTGAAGGCCAGATCGTGCTGTCAGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGCTGGTGACCAGGGCCCTGGACCGAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
 AGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCCTTCCTCTTCTCTTGGGCTTCAGACCG
 GGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC
 CTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
 CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGAAGTCTCCATCATAGAGAGCACCTGGGTGT
 CCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACCACATGGCCCAGGTA
 CACTGGAGTGGGGGTGATGTCACTATCACTGGAGAGCCATCCCCCGGGACCCCTTGAAGTGAA
 TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATTCCCATGGCGAGGACTATGCGGCCCTCTGGAGCTGCACGTGCTGGTG
 ATGGATGAGAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTGAGCAGAGGATGCAGATGCCCCCGGCTCCCCCA
 ATTCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
 CAGGTGGACCCCACTTCAGGCAGTGTGACCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
 CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCAGTGTGAAG
 TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCAGATTGGGCCCT
 ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCGCCTTCCGCCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGGGCGAAGCTGGTGGGGCCAGGCCC
 AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGTGCCACCCCCAAGTTGG
 ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTCTGCTGACCATC
 CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
 CTGCATTGAGAAATTTCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
 ACACCTACACGGTGCTTGTGGAGGGCCAGGATACAGCCCTGACTCTTGCCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT
 GGCCAGTGGGACGGTCCCTACAGCTTCACCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTACCTTGGCCCTGCATTGGGTGGAGCCAGT
 GAACACATAATCCCGTGGTGGTGCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTTCAGTGTGAT
 CGTGTGTCGCTGCAACGTGGAGGGGAGTGCATGCGCAAGGTGGGCCCATGAAGGGCATGCCCA
 CGAAGCTGTGCGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCTCATCCTC
 ATTTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
 GAAGGCGACTGTCTGAATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
 TCCCCTGGGAGAGAGCCAGCACCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
 TCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCACTT
 TATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTGTTGCCCAATAATAAAGCCCCAGAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDQASGHQATATVEVSIESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADADLEPAFRIMDFAIERGDTEGTFGLDWEPSGHVRLRLCKNLSY
EAAPSHEVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGLWCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTIVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
TCTCTTACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAATTTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCAACCACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCAACCAATTCTGAGTCCAGCACACCCTCC
AGTGGGGCCACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCAACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGTGAGGCAAGCCTGGTGGGTCCCTGGT
GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGGGCCGTGGGGCTCTTGTCTGGGC
TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
CATGGCCTCAACCATGGCCTTGCTCCAGGCCCTGGAGGGAATCATGGAGCCCCCAGAGCCAG
GTGGAGTCCCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGATTCTTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTTCTTTTCAATTCATCCAGGAGACCCTCCCAGCTTTGTTTGAGATCCTGAA
AATCTTGAAGAAGGTATTCCTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT
TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
CCCCGGGTGGGTATCTAGCTCTGAGATGAAGTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSII
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG
GGAGGCTGGAATATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTCATGGAAAGGATTTACTTTTATACTGACTCTGTTTTGGGGAAGCTTTTTTGAAGCATTT
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAGTGTATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTTCTGTGGAATTGCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGTTGGTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTATAGGAAATGGAAGGATGACAAGAGCCATTTCGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATTCAGAGAGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAACATTCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTATTCCACCTTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTACGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTGCATGACTATGTGCAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTTGCACTTAATTTTGTGGGAAAAATATTGCTACAATTTTT
TTTAATCTCTGAATGTAATTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVLLNPWSINEAVSSYCTYFIKQDSKSFQIMVSWKGIYFILTLFWGSFFGSI FMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVITGDAFVPGERSVIIMNHRTRMDWM
FLWNCMLRYSYLRLKICLKASLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDGDFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVLVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIEIACRYRLHKKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGCGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTTGTCCAGGCCCTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGTTCTTTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA
GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTTCAGTTCCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTTCCCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTC
AAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
TTTCTGATCTGAAAAGTGTAAACCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA
ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT
GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGC
TGACATGTCAGTTTGAAGGCTTGTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCCTGCTTAAAGGGC
CCCCACACCACAGACCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTGTGAAAA
CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCAGGCTCCTCATTGTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
TGAGGGCACAGTGTGCTAATGATGTGTTTTATATTATACATTTTCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTTACTTTTCTCTATACCAATCACCCATGGAATAGTTATTGAACACC
TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC
CTCATTAACACAGACACAAAAATCTAAATAAAATTTTAAACAAATTAATAAATAAATAATTTA
AAGATGATATATAACTACTCAGTGTGGTTGTCCACAAATGCAGAGTTGGTTTAAATTTAAAT
ATCAACCAGTGAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELOWRRKHGQAELOWRKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYGVCRDDVDRGKNNVTLSPNNGYWVLRILTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTLCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGGAACCCCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTTCATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC
TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
ACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
TTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
ATAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
TGACTATGATATTTCTCTGTCAGAGCTTCTAGCCCTGTTCCTACACAAATGCAGTACATAGAGTTTGTCT
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
GCCAGGGTGAATCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCCAAGCCTGGTGTATATACTAGAGTTACGGCCCTTGGGGGACTG
GATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTG
GGTGTGGAGGCCATTTTATAGATACAGAATTGGAGAAGACTTGCAAACAGCTAGATTTGACTGATCTCA
ATAAACTGTTTGCTTGATGCATGTATTTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCATCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTGTGAGAATTTGACTTGTGACATAAATTTGTAAT
GCATATATACAATTTGAAGCACTCCTTTTCTCAGTTTCTCAGCTCCTCTCATTTCAGCAAATATCCATTT
TCAAGGTGCAGAACAAAGAGTGAAAGAAAATATAAGAAGAAAAAATCCCTACATTTTATTGGCACAGAA
AAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAATGATCATATTCATTATGAAAGGTCAAGCAAGACA
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCT
ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHRLQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSSGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTI

Transmembrane domain:

amino acids 21-40 (type II)

10063710

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCTG
CCCCGATGAGCCCCGCGGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCAGGGCCCAGC
GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCTGCTCTACGCCCTCAA
TCTGCTCTTTTTGGTTAATGTCCATCAGTGTGTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTTCCCT
GTGGTTTCATCCGGTCATGATTGCTGTTTGTCTGTTTCCCTTATCATTGTGGGGATGTTAGGATATTG
TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGTGTTGCTATTTCT
GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAAGTATGGTTCCAGTACAATGGTCA
GATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGGAATTTTTTTTCAGAGAGAGTTTTAAGTGCTGTGGAGTAGTATATTTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCAGATTCCCTGCTGTGTAGAGAATTCCAGGATGTTCCAAA
CAGGCCCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTT
TTTGAGAGGAACCAAACAACTCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
ACAGACCAAATGATGTCTTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
TTGAGATGGAGGAGTTATAAAAAGAAATGTCACAGAAGAAAACCAAACTTGTTTTATTGGACT
TGTGAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
TGTTGATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTTCTACCAA
CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTTATTA
CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT
ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCACTTAAAAGAGCAAGCTAACACAT
TGCTTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAAGTCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT
ATTTTTAGCCTTTCCCTGTTAATAAAGCTTTACTATTCTGTCCTGGGCTTATATTACACATATAAC
TGTTATTTAAATACTTAACCACTAATTTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTTCAGA
AAGGACTTGTATGCTGTTTTCTCCCAAATGAAGACTCTTTTGACACTAAACACTTTTTAAAAA
GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTTCTACAGAAAA
TAGTGTTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATT
CTTTGTTTTATTTCAGTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTTTTTACAA
GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGATTTTGTGTTAT
TTCTCAGAATATGGAAAGAAAATTAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLALLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRI FEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCTCTGATAAAGCCCCCTACCAGTGCT
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTTCCAGTGTTCCTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCAACCGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCTCCTCTATGTGACAACCATGTGAATGGGAG
TGGTACCACCTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGCGCAGCGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGGAACTGCTGTCTCTGGAACACCACGGTGGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAAATGAATGTGAGCAAAACACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCGCTGT
GAGTGTGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCCCTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTGTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTTCTCTCTCAAGACATGTGGTACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGGTCTACCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
ACCTGCGAGTTTCCACGCTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCCATTCACCTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCCACCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTTGCCACCCCACTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTTCCTGCACTGCCGGGTTCTTGTCTGTGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGCGAATGCGTCTGGGGCAGGAGGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GCCCTCTAAGAACATCTGCCAACAGC
TGGGTTCAAGCTTCACACTGTGAGTTCAGACTCCAGCACCACTCACTCTGATTCTGGTCCATTCAAGTGGGCA
CAGGTCAAGCACTGCTGAACAATGTGGCCTGGGTGGGTTTCATCTTTCTAGGGTTGAAAATAAATGTCCA
CCCAGAAAGACACTCACCCATTTCCCTCATTTCTTCTTCTACACTTAAATACCTCGTGTATGGTGCAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCTAGAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTTCAATTCAAATGCAGACTAA
TTATAGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRLVLSNHTCQVPVLCKSNAIEVNIPRELVGG
LELFLTNTSCRGVSNNGTHVNILFSLKTCGTVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLVESCFATPTS KIDEVLKYILIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
 GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
 CCTGCTGCCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
 GGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
 CCCGAAGGGCCTGCAGTGGCAGTCCGCCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
 AGGGAACCTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
 GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
 GAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACGTCTTGCCTG
 TGGGAAGAGCCTGAAGACCCCCCTGTGGTGGGTGGGGAGGAGGCCCTCTGTGGATTCTTGGCCTTGGCAGG
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
 GCCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGCAG
 CTTCCCCTCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAAAGACAATGACA
 TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
 GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT
 ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGTGTGGACACCTGCCAGGGT
 GACAGTGGTGGGCCCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
 CTGCGGGGGCCCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
 GGAAGGCTGAGCTGTAATGCTGCTGCCCCCTTGCAGTGTGGGAGCCGCTTCCTTCTGCCCTGCCACCT
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
 TTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCCTCGCAGCCCAGAGGCGCCCAGAGGAAGTCA
 GCAGCCCTAGCTCGGCCACACTTGGTGTCCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
 CAGGGGTATTGCTAAGCCAAGAAGGAACTTCCACACTACTGAATGGAAGCAGGCTGTCCTGTAAAGGCC
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
 GCCTACTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
 ACTGTTGTCAATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
 AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAETACRQMGYSRAVEIGPDQDLVDVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVCSSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQGDSSGGLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAATC
TGTTTTTTGTTCTCTTGTAAGTACCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACCTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCGAAGATTTATAGGCGATGGCTCCCACTGCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCGGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTTATGTGACAGGACTTGCAATTCTCCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTTGTCCAAATTATGGGTGAGAAAAGATG
GAGGTGTTGGGTATCACAAAGGCATCGAGTCTCCTGCATTGAGTGGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTTGGCCAGGAGGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA
AACCTAAGAACCAGGTGCATTAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA
TGGTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

APP ID=10063710

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTGGGGGAGCAAGCACTTCTGGCCGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDLKNSKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
GGGCGGGCGGCCAGGATCATGTCACACCACCATGCCAAGTGGTGGCGTTCCTCCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCTT
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGCGGCTCTGTTCTGTGGGCTGGGTCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTGCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTGTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCAAAAAACAAGGAGATCCCCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACCTTTTGAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTGTCTCCCCACCCCAACTTGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAGTTACCAAAACCAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCTCTTTCTGT
CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTAATTTAAGTCCATAATATAGTTAA
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCTACATGTGGATAGAAG
GAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC
AGCTGGCTAAAACATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTCTTG
GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTTTGGAACTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTGCCAATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCCAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAGTTATCACACCGAAAGAAGTCAACGAGCG
TCTACTCCAGAAGTCAGTATGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
ACAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAATGAGATA
TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTTATCATTTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGAACATTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCA
TTTATAATGAAGATTAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAAGTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
CTTGTGTATTAAATTAACATTTTTAAAACGCAGATATTTGTCAAGGGGCTTTCATTCAAACGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
AAAATATTTTTGTTTTGTATTTGAAGAAGATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATA
TGCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTTGTTCTGTGAAAAATAAATTCCTTCTTGTA
CCATTTCTGTTTAGTTTTACTAAATCTGTAAATACTGTATTTTCTGTTTATCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTCGACGTCTGTCTAGCTTAAATGAATGTGTCTATTTGCTT
TATACATTTATATTAATAAATTTGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMQPWRVSAFIENNIVVFENFWGLWMNCVRQANIRMQCK
IYDSSLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCCCTCTCCTCGGTCTCCTCCGCCCTCCAGCTCCGCGCTGCCCGGCAGCC
GGGAGCCATGCACCCCAGGGCCCCGCGCCTCCCCGAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTGCGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACCTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTCACTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTCAAGGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATAAAATGCTTTAATTTTCATTTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACCTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTCTTTCTCTTAGTATAGCATTTTAA
AAAAAATATAAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKEGEGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACCTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCGTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCAAGGGGTCCAATTTT
TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
GCCCCAAGCCAAAGCAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
ATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAATGGAATACGCAGACTCAAAGA
GCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATT
TACGGAACCTGGATCTGTCTTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
TCTTTGCTGGCATGATCAGACTCAAAGAACCTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
CTGGCCCTTTTTCCAAGGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGAATAAAATCAGTGT
CATAGGACAGACCATGTCTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
TCGAAGCTTTTCAGTGGACCCAGTGTTTTCCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT
TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGATTCTTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG
GGCTCTCCCAAAGCCGACGTTTAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAAACCCCTT
TGCCCCCGACGGTGGGAGCCACAGAGCCCGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCTCTGCTGGTTAT
CTACGTGTCTGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA
GATTATAAACCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTCTTAAAGCT
GGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCTCCCC
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCTTGTCGGTTTTAGTGCATTATAATACT
GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
GAACCTCCGGTTTAAATATAATACCTATTGTATAAGACCTTTACTGATTCCATTAATGTCGCATTT
GTTTTAAGATAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTPLLTLSSAERGCPKGCRCGKMVYCESQKLQEIPSSISAGCLG
LSLRYNLSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRITPVRIQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFCVFNLRNLDSNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAHEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKRQSLKQMTPTQEFYVDYKPTNTETSEMLLNGTGPCCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCCTCCGTCCTGGAGGAGGAGGAGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT
AGCAACCTTTCCCTGGATCTCACAAAACCTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTTGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAACCTGACCTAGTGAAGGTCAGATGCAATGGAAGGAAAAAGGAACTGGAAGG
AAAACCATTGCGATTTCTGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTTACC
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFVGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCCTCGGGGGGCATCGGCGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA
CCCCTGCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCTTCATGGCTTGCCCTCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT
CACGGGATACCACTTCCCTGTCCACACCCCGACCAGGGGCTAGAAAATTGTTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTTACTTGTTAACTTGTTCTTGTTGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCTG
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCCAGTCTTGGCTTCTTGTCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCCTGGCCCACTGGATTTCATGGTGATCATTAAAAAGAAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQQLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSVDICINNAGLARPDTLTLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPVTHFYATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AAC TTCTAC ATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTT CAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAAGAATCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAAC TTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTT CATTTCAGTAGCCACCAGCCACCTGTGG
CCGTGAGTGCTTGAAAT GAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA
TTAATTTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCAACTAATTTTTGTATTTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTTCTTTGTGTGGGAAC TTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTGTTCTTAATTTCAATTAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGGHTLGESPMPEKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
 ATCTTCCTCATCGGGACTAAAATGGGCTGTTCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
 TCCATCTGTGTGCTGCGTACGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG
 GAATACCAGAGGATGCTACAACCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCTTCAGAT
 TTGAAAACTTGCTGAAAGTAGAAGAATATACCTATAACCACAACAGTTTAGATGAATTTCTACCAACCT
 CCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACCTTCAA
 AAATTCCTATCTGGAAGAATTACATTTAGATGACAACCTGTCTCTGCACTTAGCATAGAAGAGGGAGCA
 TTCCGAGACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAATTCCTGCGGTTT
 GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTCAAG
 GTCTCACTAGTCTAAAACGCTGTTCTAGATGGAACTGTTGAACAATCATGGTTTAGGTGACAAAGTT
 TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAAACCT
 TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTTT
 CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
 GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
 GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
 AGGTTCTGTTGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
 ACCATTGAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCAAGGACAGTGGCCAGCTCCAGTGAC
 CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
 CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTATCTCTTGAAACTTGCTCTACCTATG
 ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGT
 AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCCTATAAAGTATGCATGGTTC
 CCATGGAAACCAGCAACCTCTACCTATTTGATGAACTCCTGTTTGATTGAGACTGAACTGCACCCCTT
 CGAATGTACAACCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
 TTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGTATG
 TTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
 GCAGAAGCTGGCACTAAGAAGGACAACCTATCTGGAATCAGGGAACTTCTTTTCAGATGTTACCAAT
 AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAACACACCATATTTCTCCTAATGGAATGAATCTGT
 ACAAAAACAATCAGAGTAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
 TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
 GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIR TITYDSLSKIPYL
EELHLDNDSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRLQLYRLDMSNNNLSNLPQGI FDDL DNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQQQWPAPVTKQPD
IKNPKLTKDQQTTGSPSRKTITITVKS VTS DTHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVMVPMETSNLYL FDETPVCIETETAPLRMYNPTTTLNREQEKE
PYKNPNPLAAIIGGAVALVTIALLALVCWVYHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNL YKNNHSESSSNRSYRDSGIPDS DSHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCAGGCCACCCAGGCTTCTTGGA
GCCCTGCCGGGCCACTTGTCTTCATGCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTTCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCAGGAGCCCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCTCCAGCCCGG
GTTTTGGAAGCTCCCACATGCCTGGATCCCACTGATGCCTCCTTGGTGTACCCACGTTCTGGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGTGTGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCCACCACTGCTCTTCTCCTCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCTACCCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
TGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAGCTATTCAA
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
AATTCTCTTGCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCCAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGTTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLGLARAQGAERTVSSEQRPAMASLGLLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSYLHKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAATTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCGAGTGTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC
AGCTTACAGATTTGATCCC GTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTGCACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAAAGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCCCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGFLKTDGSFVVDIPSGSYVVEVVSAPAYRFDVPRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATCCCCAAAACAA
GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTGGTCTCCAGATTTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTGCTAAGACTCTATCTGGACAGGGTATTTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTA
ATGTATTTATTTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATTTTATAACCTG
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAATAAACCTTGTAATCTAGAAAGAGTGG
CTAGGGGGGTATTTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTGTGATGTGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
CCAGGAATCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTCTTTCATACCAAAAAAAA
AAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPG LTCWALTAEPGWGQNGATT CATNSHSDSELRPEIF
SSREAWQFFLLWSPDFRPKM KASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLT IKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLOWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTTCGGTTCCGCATCCAGCC
TAGCGTGTCCACGATCGGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCTTCGGGGATTCTCCCGGCTCCCGTTCGTTCCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
TTCCCAAAGCATTTTGTGGAATATGATGGAACAACCTCATTTTTCGTGTCAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTCAGGGCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCCCTTTACCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAG
AATGTGCCGTGATATGAAAAGATCCTGGGTTTGAAGCAGTTTAAATGTCAGAAAGATTGCATGG
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTTCAGAACTCCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCATCTCCTGGGTTTTCTCTGCTCTTTTATTGGTGATCCTGGTTCTTTCGGCCGTTTACGT
CATTGTGTGCACCTCAGCTGAAAGTTCTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT
GCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCCTGGAC
AGGGGGCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCCTCTGGGTGTCCCAGACAGGTG
TTCACATCTGTGCTGTGAGGTGAGTGCCTCAGTTCTTGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGTAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCAAGATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTTGGAATTCTTAGTCCTTGGCC
TCGGACACCTTCATTTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCAGGATCAAGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
ACCCCAACCTGACAGCCCTCATCCCTCTTGGCTTGAAGCGTCAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTTGTGATCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG
CCACGCTTGACCTCGGGCCCATCTGGGCTCATGCTCTCTCCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGHGAEPPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCAGATAAAAACATACGCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTTGAACTAATTGTATAAAAAACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLDSGNLIAVPDKNYIRPEI
FFALASSLSSASAEKGSPIILGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCCTTCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACCTGCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGAGGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTTTGTATCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAACATCAGCATTTTAAGAAAAAAAAAA
AA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCTGGCTTGTCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMLVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATT
CCTGATGATTTATAGACTCAAAGAAAACTCATGTTTCAGAAGCTCTCTTCTCTTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTATTTTTTGAAATTT
CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSECKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAAGTTTACAGCACCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCGCGCGCCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCGCCCCAGTCC
CGGCCCCCTCTCCCGCCCCACACCCACCCTCCTGGCTCTTCTGTTTTACTCCTCCTTTTCATTCATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGGAAGAATGGGGTT
CCTCGGGACCGGCACTTGGAATCTGGTGTAGTGCTCCCGATTCAAGCTTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG
AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCCACAA
TTATGAGGAGGATCCCAATAAGCCCAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACCTCA
ATATTTCCCAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAGGAATATGGAAGCTTGAAGGATTCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAAATTTTTGACCAAGG
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDDYSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVT PMAAIQDGLAKGENDET VSNLTLTNGLE
RRTKTYS E DNF EELQYFPNFYALLKSIDSEKEAKEKETLITIMKT LIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLF PAPS EKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIAKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAGGGTGGAAG
CCAGTGCCTGTATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCAGTCTCTCTGCTCTCAGGACCCACGCTGACTTAG
TGGGCACCTGACCACTTTGTCTTCTGGTTCCAGTTTGGATAAAATTCTGAGATTTGGAGCTCAGT
CCACGGTCTTCCCCACTGGATGGTGCTACTGCTGTGGAACCTTGTA AAAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAAGATTTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT
TAATGGTAAGTACAAGTGTTACCTGAGCCCCGAGGCCAACCATCCCCAGTTGAGCTTATA
GGGTGAGTAGCTCTCCACATGAAGTCTGTCACTCACCAGTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGCCCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCTATA
TGCTAGCTTTCCCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGATGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA
GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTATTTTAAATTAAGATACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT
AGGTGATTTTTCTTTAATTCTGTTAATTTATCTGTATTTCTTAATTTTCTACAATGAAGATGA
ATTCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAGGAACCTAAAGAACA AAAATCATCTGGTAATCTTTCTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAACTCTCAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTTCCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAAAGAGGCAAGAAGAATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAATAAATTTCCGGCTGTTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
CTCCTCGTGTTTACATTCTGTGTGTGTCCCCCTCCACAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTGCCCCACCGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTGGTCTAACTTGTTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESVVPNRWLDASLSPVILGVQGGG
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPIITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTT
CAGCCTTATATGCAGGAGGTGGTGGCCCTTCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAAGTGGATTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTGTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTTT
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTATATCATTTTATTAATATGGATTATTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCCGCATGTACCAGGTGAGTGCAGAGGGC
TGCCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCTACAGGTGGTTGCAT
TCTTGGAATGGTCATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCCGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCACGCCCCT
TGCCCTGTGCCCCGACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGCGGCCCCGT
GTGATGGGCTAGCCCGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTCTGTGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTCTGTCAATTTCTCTCAGGAAAGGTTTCAA
GTTCTGCCCATTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCTG
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCCTGGTTTTATT
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHYSHWPSCCPKQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTTCATTGGGGCCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGTAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGAAGTGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCTGCTGTCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCATTAAAGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCAATTGCAGAAGTGAGGTATCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCA
GCAGACAAAGTCGTCTTCTTCTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAAGTCTCAAGACCTCTTCCCCCTTGCCCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTGAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLDLRLRVEPVTTSVATGDYSILMNVS
WRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCACTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATCCGTTCCCATCCAGCAA
GAGACCCTGGTCGTCGGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT
GGTGA CTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKT LHGPAMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSM SRNIESRSTSPWNYT VTWDPNRY PSEVVQAQCRNLGCINAQ GKEDISMNSVPIQQETLVV
RRKHQGC SVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

A C A C T G G C C A A A C A A A A A C G A A A G C A C T C C G T G C T G G A A G T A G G A G G A G A G T C A G G A C T C C C A G G
 A C A G A G A G T G C A C A A A C T A C C C A G C A C A G C C C C C T C C G C C C C C T C T G G A G G C T G A A G A G G G A T T C
 C A G C C C C T G C C A C C C A G A C A C G G G C T G A C T G G G G T G T C T G C C C C C C T T G G G G G G G G C A G C A C
 A G G G C C T C A G G C C T G G G T G C C A C C T G G C A C C T A G A A G A T G C C T G T G C C C T G G T T C T T G C T G T C C T
 T G G C A C T G G G C C G A A G C C C A G T G G T C C T T T C T C T G G A G A G G C T T G T G G G C C T C A G A C G C T A C C
 C A C T G C T C T C C G G C C T C T C C T G C C G C C T C T G G G A C A G T G A C A T A C T C T G C C T G C C T G G G G A C A T
 C G T G C C T G C T C C G G G C C C C G T G C T G G C G C C T A C G C A C C T G C A G A C A G A G C T G G T G C T G A G G T G C C
 A G A A G G A G A C C G A C T G T G A C C T C T G T C T G C G T G T G G C T G T C C A C T T G G C C G T G C A T G G G C A C T G G
 G A A G A G C C T G A A G A T G A G G A A A G T T T G G A G G A G C A G C T G A C T C A G G G G T G G A G G A G C C T A G G A A
 T G C C T C T C C C A G G C C C A A G T C G T G C T C T C C T T C C A G G C C T A C C C T A C T G C C C G C T G C G T C C T G C
 T G G A G G T G C A A G T G C C T G C T G C C C T T G T G C A G T T T G G T C A G T C T G T G G G C T C T G T G G T A T A T G A C
 T G C T T C G A G G C T G C C C T A G G G A G T G A G G T A C G A A T C T G G T C C T A T A C T C A G C C C A G G T A C G A G A A
 G G A A C T C A A C C A C A C A C A G C A G C T G C C T G C C C T G C C C T G G C T C A A C G T G T C A G C A G A T G G T G A C A
 A C G T G C A T C T G G T T C T G A A T G T C T C T G A G G A G C A G C A C T T C G G C C T C T C C C T G T A C T G G A A T C A G
 G T C C A G G G C C C C C A A A A C C C G G T G G C A C A A A A C C T G A C T G G A C C G C A G A T C A T T A C C T T G A A
 C C A C A C A G A C C T G G T T C C C T G C C T C T G T A T T C A G G T G T G G C C T C T G G A A C C T G A C T C C G T T A G G A
 C G A A C A T C T G C C C C T T C A G G G A G G A C C C C G C G C A C A C C A G A A C C T C T G G C A A G C C G C C C G A C T G
 C G A C T G C T G A C C C T G C A G A G C T G G C T G C T G G A C G C A C C G T G C T C G C T G C C C G C A G A A G C G G C A C T
 G T G C T G G C G G G C T C C G G G T G G G G A C C C C T G C C A G C C A C T G G T C C C A C C G C T T T C C T G G G A G A A C G
 T C A C T G T G G A C A A G G T T C T C G A G T T C C C A T T G C T G A A A G G C C A C C C T A A C C T C T G T G T T C A G G T G
 A A C A G C T C G G A G A A G C T G C A G C T G C A G A G T G C T T G T G G G C T G A C T C C C T G G G G C C C T C T A A A G A
 C G A T G T G C T A C T G T T G G A G A C A C G A G G C C C C A G G A C A A C A G A T C C C T C T G T G C C T T G G A A C C C A
 G T G G C T G T A C T T C A C T A C C C A G C A A A G C C T C C A C G A G G G C A G C T C G C C T T G G A G A G T A C T T A C T A
 C A A G A C C T G C A G T C A G G C C A G T G T C T G C A G C T A T G G G A C G A T G A C T T G G G A G C G C T A T G G G C C T G
 C C C C A T G G A C A A A T A C A T C C A C A A G C G C T G G G C C C T C G T G T G G C T G G C C T G C C T A C T C T T T G C C G
 C T G C G C T T T C C C T C A T C C T C C T T C T C A A A A A G G A T C A C G C G A A A G G G T G G C T G A G G C T C T T G A A A
 C A G G A C G T C C G C T C G G G G C G G C C C A G G G C C G C G C G C T C T G C T C C T A C T C A G C C G A T G A
 C T C G G G T T T C G A G C G C C T G G T G G G C G C C C T G G C G T C G G C C C T G T G C C A G C T G C C G C T G C G C G T G G
 C C G T A G A C C T G T G G A G C C G T C G T G A A C T G A G C G C G C A G G G C C C G T G G C T T G G T T T C A G C G C A G
 C G G C G C C A G A C C C T G C A G G A G G G C G G C G T G G T G T C T T G C T C T C T C C C G G T G C G G T G G C G C T
 G T G C A G C G A G T G G C T A C A G G A T G G G G T G T C C G G G C C G G G C G C A C G G C C C G C A C G A C G C C T T C C
 G C G C C T C G C T C A G C T G C G T G C T G C C C G A C T T C T T G C A G G G C C G G G C G C C C G G C A G C T A C G T G G G G
 G C C T G C T T C G A C A G G C T G C T C C A C C C G A C G C C G T A C C C G C C T T T T C C G A C C G T G C C G C T C T T
 C A C A C T G C C C T C C C A A C T G C C A G A C T T C C T G G G G G C C C T G C A G C A G C C T C G C G C C C C G C G T T C C G
 G G C G G C T C C A A G A G A G A G C G G A G C A A G T G T C C C G G G C C C T T C A G C C A G C C C T G G A T A G C T A C T T C
 C A T C C C C C G G G G A C T C C C G C G C C G G A C G C G G G T G G G A C C A G G G C G G G A C C T G G G G C G G G G A
 C G G G A C T T A A T A A A A G G C A G A C G C T G T T T T T C T A A A A A A

FIGURE 162

MPVPWFLLSLALGRSPVVLSELERLVGPQDATHCSPGLSCRLWSDIILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSFQAYPTARCVLLEV
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVVSADGDNVHLVLNV
EEQHFGLSLYWNQVQGPVKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQ
LWQAARLRLTLQSWLLDAPCSLPAAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ
VNSSEKLQEQECLWADSLGPLKDDVLLLETRGPQDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSAAAARG
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLFSP
GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLHHPDAVPALFRTVPVFT
LPSQLPDFLQALQQPRAPRSGRQLQERAEQVSRAALQPALDSYFHPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCACTTTGA
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTACGCGGATCACCCGGAAGTCTGTC
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
 GTCAGTGCGGGAGGCGGTCAGCCACCAAGATGACTGACAGGTTTCTGCTGTCAGCACACTAC
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGATGATTGTTTCATCCTA
 CCCCCACGCCAATCCGTGACGGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCATGACCTG
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA
 ATATGAGTTCTTCGGCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTGCGTTCCCA
 CCTGGGCCAAGGAGAGTGGCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCTTCTCCGGAGCCTTCTGTCTCCATGGGCTTCTCGTGCAGTACTCTGCTACCTGAG
 CTACAGATATGTCAACAGCCGCTGACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
 TCCAGCCGCTGCGCTTCATCCAGGAGCAGTCTGATCCCTGCTTTGACCTCAGCGGCCCCAGC
 AGTCTGGCCCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCCGAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCCAAACGCTGCCCTGAG
 GTCGGGCCCCCATCTATGCACCTCAGGTGACCCCGAAGCTCAATTCCCATTCTACGCCCCACA
 GGCCATCTCTAAGGTCCAGCCTTCTCTATGCCCCCAAGCCACTCCGGACAGCTGGCCTCCCT
 CCTATGGGGTATGCATGGAAGTTCTGGCAAAGACTCCCCCACTGGGACACTTTCTAGTCTTAA
 CACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
 TTCTCTGACGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTCACCAGC
 CCCTGGGGATTTGCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGAC
 CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTTGTGTGCCAAGGATGAAGCCAAGAGCCAGCCCCCTGAGACCTCAGACCTG
 GAGCAGCCACAGAACTGGATTCTCTTTTTCAGAGGCTGGCCCTGACTGTGAGTGGGAGTCC**TG**
AGGGGAATGGGAAGGCTTGGTGCTTCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGCTCGCTGAGCCCTG
CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATCCTGGCCAGTTTC
ACAATCTAGCTCGACAGAGCATGAGGCCCCCTGCCTCTTCTGTCTATTGTTCAAAGGTGGGAAGAGA
GCCTGGAAAAGAACCAGGCTTGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGC
ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGTCA
TTCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTACAGCTTCATTCTCTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCCT
ATCCTGAGAATGGGGTTTGAAAGGAAGGTGAGGGCTGTGGCCCCCTGGACGGGTACAATAACACAC
TGTACTGATGTCAACCTTTGCAAGCTCTGCCTTGGGTTACGCCCATCTGGGCTCAAATTCAGC
CTCACCCTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCCGAAGCTCGGTTTCTCT
ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTTAAAGTGCTTAATAGTGCTTGTACATGGGCAGTGCCCAATAAACGGTAGCTATTTAAAAA
 AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTELYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCTIS
KVRSIQMIHVHTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRITYQMHGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLESMGFLVAVLCYLSYRYVTKPPAP
PNSLVQVRVLTTFQPLRFIQEHVLPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPSPSYAPQVTPEAQFFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPGQLQKEPPAGSCMLGGLSLQEVTSIAM
EESQEAQSLHQPLGICTDRTSDPNVLHSGEETPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESIVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTCACCCGGGCGCGGTGGCCACAACATGG
CTGCGGCGCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTTGGGCTGGA
AGTGTGTAACACAGTTTTGGATATTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCCTGTGCCTGAGCCCGAGGCATTTCAGAGCTGATTTCAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTTGTTGGCATCAAAATCTAAGTTTGTTTTACAAAGATTGTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAAATATATTATTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFLVGLALWWVPGQSDLHGRRFSDLKVCGDEECSMLMYRGKALEDFTGPD CRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPAETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPVEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRADSEDGEGA
FSESTGLQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL SLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCGGCAGGGAGGAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCCCGCGCGTGGTACAG
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAACT
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCTCCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
TGCCTGCCCAGATCTCTACCAGTGGTCTGATGGAAGCAATCCCAGTACCGAACTGGTACACAG
ATGAACCTTCCTGCGGAAGTGAAAAGTGTGTTGTGATGTATCACCAACCAACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTTACCAGTGGAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACCTGTTGTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCCAGAATTTTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGGAAATGGCTTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAAGTAATTTTTATATGTCTATTATTTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
ACTTCAAACCTTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTACCTTTTATAAGTTGTTATCTAGTCAATGTAA
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGGTTGATAAAA
ATGAACTGTTCTAATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
AATAAGAAGCTATTTTCATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTATAGCTTAAAATTAAACAGATT
TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAAG
TGACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTCGGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAAATGGGTTGG
AACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGCTTGAGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTCTTCTCAAGAGAAAGTTGTAACCTCT
CTGGTCTTCATATGTCCCTGTGCTCCTTTAACCATAAAGAGTCTTGTTTCTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFRRVVSQGKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSCKGRKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217